

James Mayfield

TAC 2013 Cold Start Knowledge Base Population

Special thanks to:
Hoa Trang Dang
John Frank
Ralph Grishman



The Task

You are given:

Schema

```
per:children  
per:other_family  
per:parents  
per:siblings  
per:spouse  
per:employee_of  
per:member_of  
per:schools_attend
```

When Lisa's mother Marge Simpson went to a weekend getaway at Rancho Relaxo, the movie The Happy Little Elves Meet Fuzzy Snuggleduck was one of the R-rated european adult movies available on their cable channels.

Schema

per:children
per:other_family
per:parents
per:siblings
per:spouse
per:employee_of
per:member_of
per:schools_attend

You Must Produce:



per:schools_attended



per:children

per:cities_of_residence

per:spouse



per:children

BOTTOMLESS PETE, NATURE'S CRUELEST MISTAKE



per:alternate_names

per:alternate_names

When **Lisa**'s mother **Marge Simpson** went to a weekend getaway at Rancho Relaxo, the movie **The Happy Little Elves Meet Fuzzy Snuggleduck** was one of the R-rated european adult movies available on their cable channels.

How do you know
that your KB is any good?

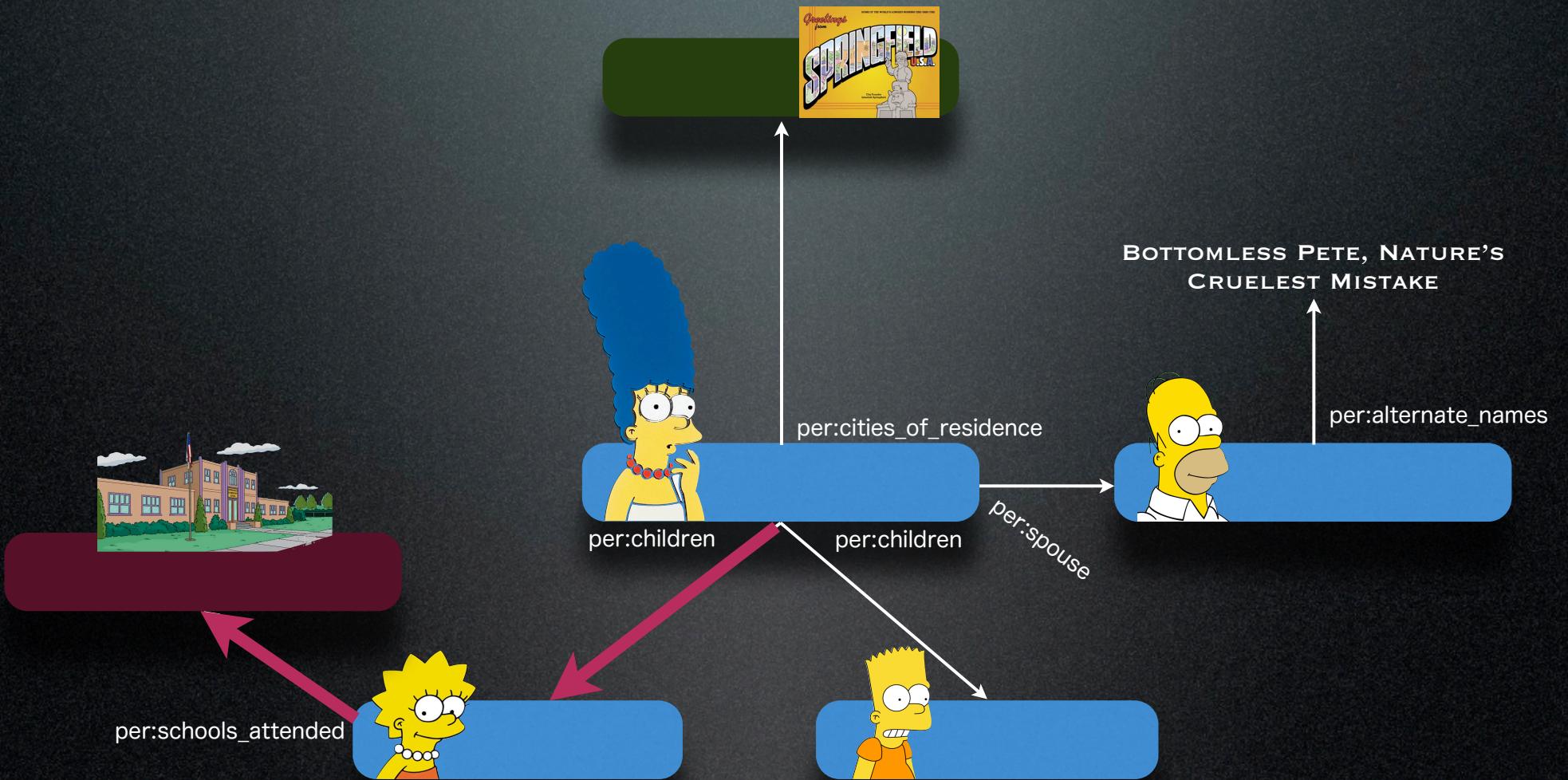


Align it to a
ground
truth KB



But how are you
going to produce ground
truth? And wouldn't the
alignment be intractable
anyway if the KB were of
any reasonable size?

Where did the children of Marge Simpson go to school? per:children
per:schools_attended



When Lisa's
mother
Marge
Simpson
went to a
weekend
getaway at



After two
years in the
academic
quagmire of
Springfield
Elementary,

**BOTTOMLESS PETE, NATURE'S
CRUELEST MISTAKE**



per:schools_attended



per:children

per:cities_of_residence

per:children

per:spouse



per:alternate_names

Entity-Valued Relations

Relation

per:children
per:other_family
per:parents
per:siblings
per:spouse
per:employee_or_member_of
per:schools_attended
per:city_of_birth
per:stateorprovince_of_birth
per:country_of_birth
per:cities_of_residence
per:statesorprovinces_of_residence
per:countries_of_residence
per:city_of_death
per:stateorprovince_of_death
per:country_of_death
org:shareholders
org:founded_by
org:top_members_employees
{org,gpe}:member_of
org:members
org:parents
org:subsidiaries
org:city_of_headquarters
org:stateorprovince_of_headquarters
org:country_of_headquarters
org:country_of_headquarters

Inverse(s)

per:parents
per:other_family
per:children
per:siblings
per:spouse
{org,gpe}:employees_or_members*
org:students*
gpe:births_in_city*
gpe:births_in_stateorprovince*
gpe:births_in_country*
gpe:residents_of_city*
gpe:residents_of_stateorprovince
gpe:residents_of_country*
gpe:deaths_in_city*
gpe:deaths_in_stateorprovince*
gpe:deaths_in_country*
{per,org,gpe}:holds_shares_in*
{per,org,gpe}:organizations_founded*
per:top_member_employee_of*
org:members
{org,gpe}:member_of
{org,gpe}:subsidiaries
org:parents
gpe:headquarters_in_city*
gpe:headquarters_in_stateorprovince*
gpe:headquarters_in_country*
gpe:headquarters_in_country*

String-Filled Relations

per:alternate_names

org:alternate_names

per:date_of_birth

org:political_religious_affiliation

per:age

org:number_of_employees_members

per:origin

org:date_founded

per:date_of_death

org:date_dissolved

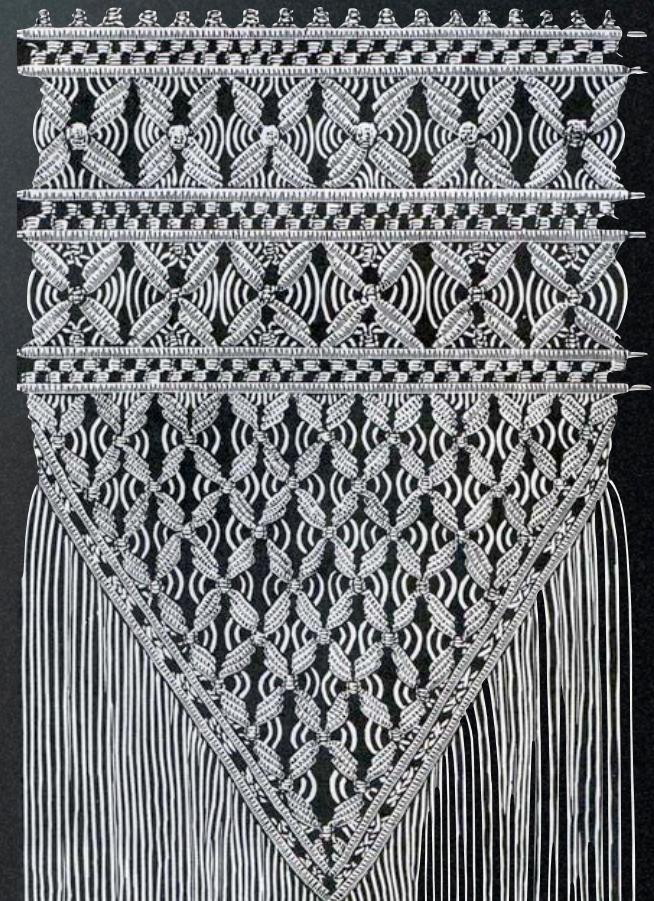
per:cause_of_death

org:website

per:title

per:religion

per:charges



Sample Evaluation Queries

Query Entity	First Relation	Second Relation
Appleton Museum of Art	org:top_members_employer	per:title
Annie Laurie Gaylor	per:employee_or_member_of	org:stateorprovince_of_headquarters
William Shore	per:organizations_founded	org:date_founded
Wistar Institute	org:employees	per:title
Andrew W. Mellon	per:children	per:organizations_founded
Lycee Alliance Israelite Universelle	org:employees	per:schools_attended
Tsitsi Jaji	per:schools_attended	org:students



2013 Document Collection

50K TREC KBA Documents drawn from two localities: Ocala Florida and Central Kentucky

#Docs	Domain
10139	chronicleonline.com
12089	ocala.com
159	ocaladailyphoto.blogspot.com
1840	thetimestribune.com
675	wruf.com
4104	www.centralkynews.com
20959	www.kentucky.com
49965	Total



Sample Cold Start Document

Men in Winchester bingo theft receive sentence -
centralknews.com

The men, of paying volunteers to work charitable gaming events in the Winchester Youth Baseball League pleaded not guilty this morning.

Larry K. Roe, 47, of 2080 Wood Pointe, and Mark F. Thomas, 49, of 140 Casa Loma made their first appearance in Clark District Court. Both men were arrested in October after state Department of Charitable Gaming officials obtained warrants against both for theft.

Since the charges are felonies, the judge entered an automatic not guilty plea for both men and scheduled a preliminary hearing for Dec. 14.

Hah! You Believed Me.

Sample Cold Start Document (for real)

<DOC id="1320167760-c04e70146438f77b643e8440c009a0f6">
<SOURCE><http://www.centralkynews.com/winchestersun/news/ws-men-plead-not-guilty-in-bingo-theft-case-20111101,0,3013153.story></SOURCE>
<TEXT>

Men in Winchester bingo theft receive sentence - centralkynews.com

FindKyJobs

e-Edition

Today's Print Ads

Classifieds

Post an Ad

GetItDoneKy.com

Pet Place

Joe Ellis



Submitted KB Sizes



	NYU1	UMass_IESL1	hltcoe1	hltcoe2	hltcoe3	hltcoe4	hltcoe5
PER	108575	116771	121934	113159	112496	111718	119940
ORG	40545	78184	66154	62179	61887	60472	65765
GPE	16311	19172	13141	13640	13548	13497	14316
Mentions	2410664	2640975	3701152	3551434	3471084	3266734	3828903
Canonical Mentions	1199204	1425309	2100715	2004733	1991744	2049164	2173286
Slots	67813	475917	221019	209835	251287	222534	326662

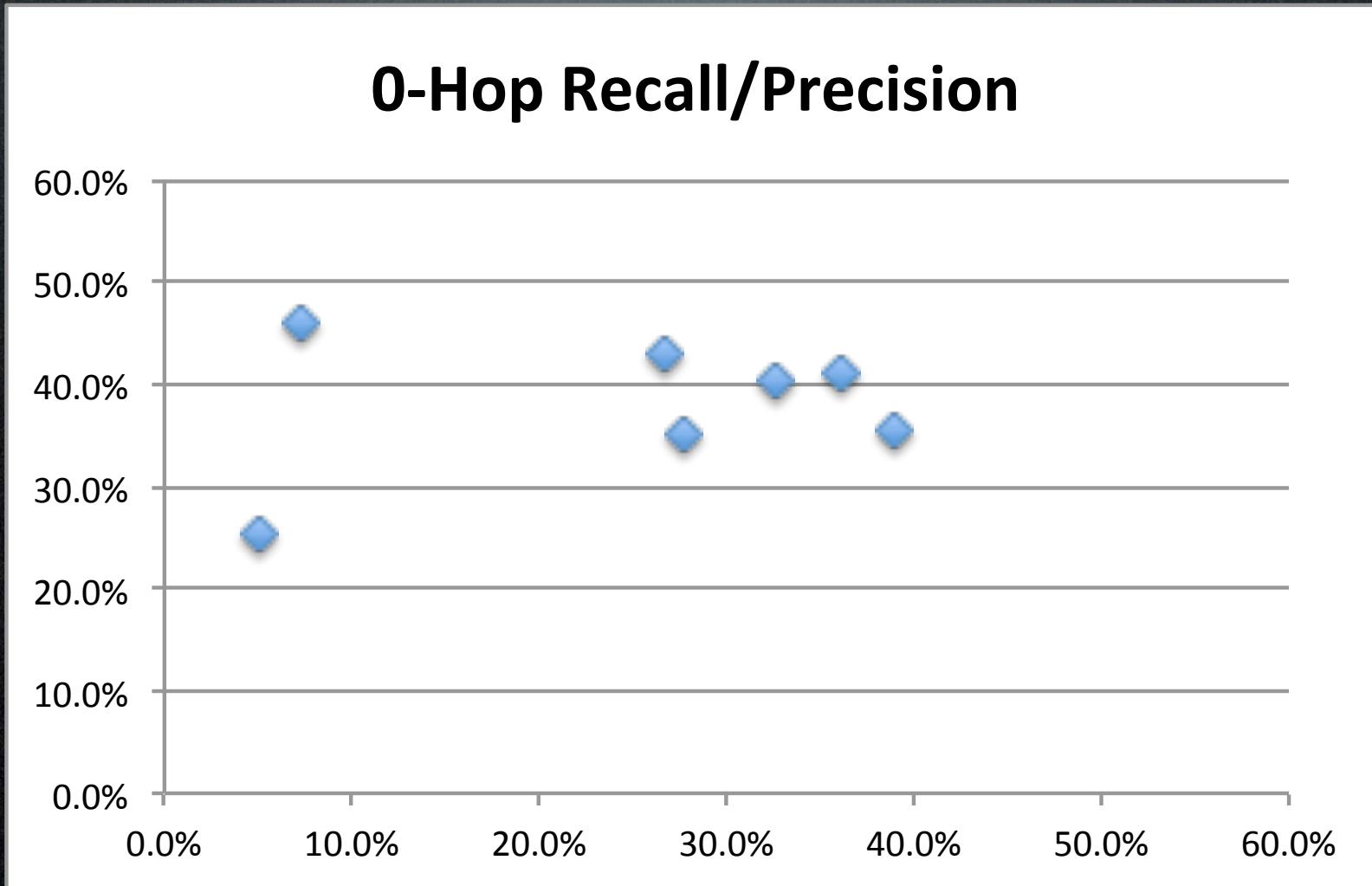
Submitted Relation Counts

	NYU1	UMass_IE	hltcoe1	hltcoe2	hltcoe3	hltcoe4	hltcoe5
gpe:births_in_city	200	564	95	87	81	86	110
gpe:births_in_country	25	1304	2203	2021	1984	2026	2511
gpe:births_in_stateorprovince	41	353	36	32	86	89	128
gpe:deaths_in_city	57	71	164	137	129	128	181
gpe:deaths_in_country	8	56	62	57	87	86	155
gpe:deaths_in_stateorprovince	13	589	55	72	163	151	227
gpe:employees_or_members	1090	0	6840	6410	6350	5607	7346
gpe:headquarters_in_city	154	1466	4354	4059	3929	3910	4863
gpe:headquarters_in_country	43	988	1144	1096	1915	1862	3438
gpe:headquarters_in_stateorprovince	44	5418	999	1022	3270	3194	4502
gpe:member_of	0	0	101	87	93	83	110
gpe:organizations_founded	0	0	9	10	10	8	12
gpe:residents_of_city	1487	832	6681	6464	6200	6209	7943
gpe:residents_of_country	86	1076	4722	4314	5894	5667	10954
gpe:residents_of_stateorprovince	338	1328	2585	2626	14259	12619	30309
gpe:subsidiaries	0	0	1684	1809	1743	1423	2039
org:alternate_names	15111	109319	39322	37129	36916	25062	39432
org:city_of_headquarters	150	1411	4354	4059	3929	3910	4863

	Count	Nodes	Edges	Count	Nodes	Count	Nodes
org:city_of_headquarters	150	1411	4354	4059	3929	3910	4863
org:country_of_headquarters	42	931	1144	1096	1915	1862	3438
org:date_dissolved	2	0	117	113	114	105	118
org:date_founded	82	431	243	243	243	208	261
org:employees_or_members	2593	5589	33697	31658	32230	29240	36993
org:founded_by	62	2955	752	722	723	669	839
org:member_of	35	141	219	222	202	178	253
org:members	35	141	320	309	295	261	363
org:number_of_employees_members	9	0	0	0	0	0	0
org:organizations Founded	0	0	70	70	74	67	86
org:parents	115	901	4455	4481	4382	3687	5051
org:political_religious_affiliation	1	0	0	0	0	0	0
org:shareholders	0	115	19	21	21	16	20
org:stateorprovince_of_headquarters	42	5101	999	1022	3270	3194	4502
org:students	185	228	1124	1108	1106	906	1285
org:subsidiaries	115	901	2771	2672	2639	2264	3012
org:top_members_employees	205	4277	465	431	4906	4494	5644
per:age	10328	8887	1657	1623	1608	1558	1701
per:alternate_names	283	162318	4346	4231	4196	3154	4480
per:cause_of_death	218	2167	0	0	0	0	0
per:charges	18	8981	843	825	815	750	870
per:children	876	1615	1297	1246	1228	1083	1390
per:cities_of_residence	1487	832	6681	6464	6200	6209	7943
per:city_of_birth	200	521	95	87	81	86	110
per:city_of_death	57	65	164	137	129	128	181
per:countries_of_residence	86	1076	4722	4314	5894	5667	10954
per:country_of_birth	25	1246	2203	2021	1984	2026	2511

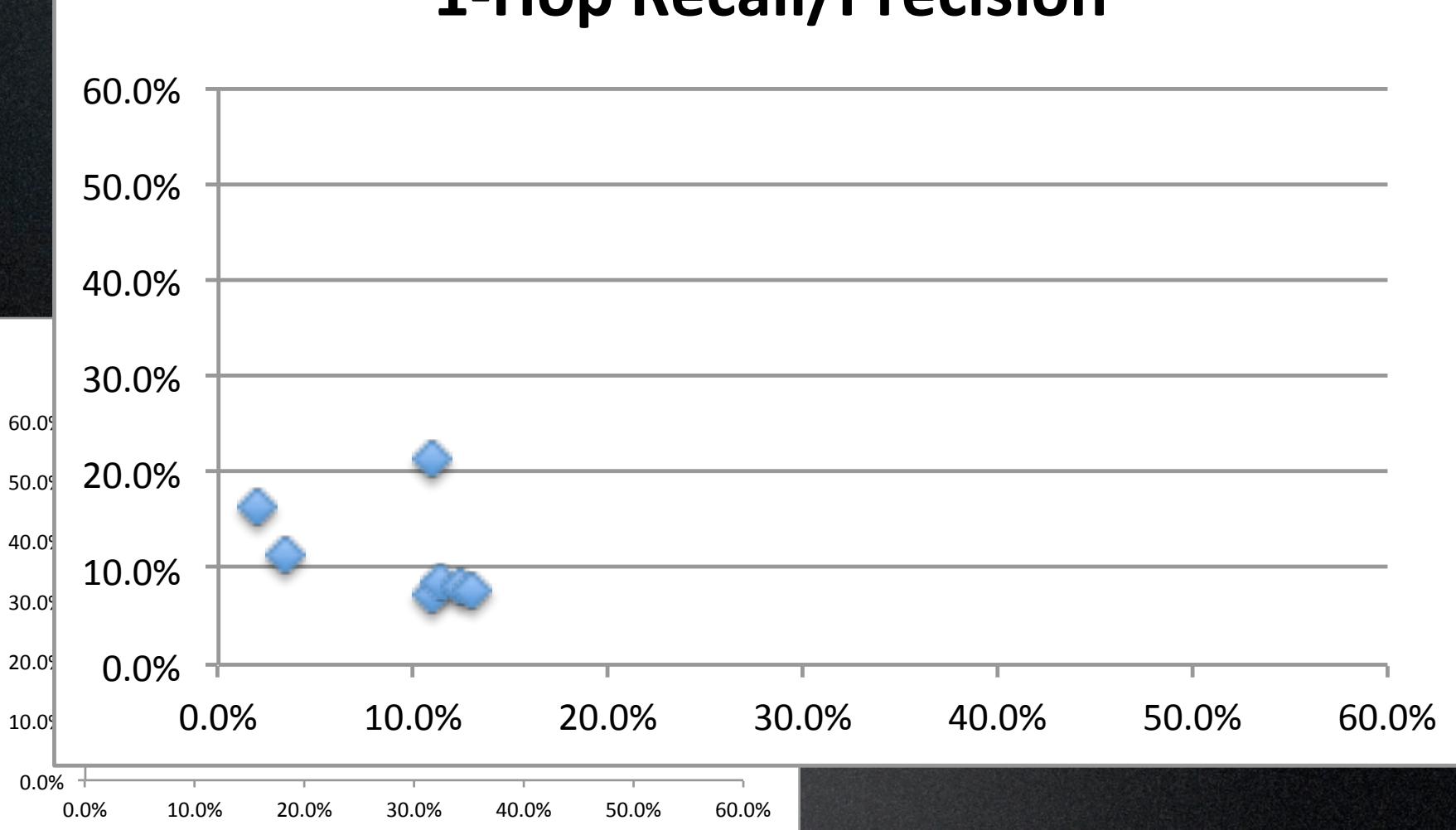
per:city_of_birth	200	521	95	87	81	86	110
per:city_of_death	57	65	164	137	129	128	181
per:countries_of_residence	86	1076	4722	4314	5894	5667	10954
per:country_of_birth	25	1246	2203	2021	1984	2026	2511
per:country_of_death	8	56	62	57	87	86	155
per:date_of_birth	64	420	291	285	281	236	301
per:date_of_death	883	0	1834	1748	1736	1591	1828
per:employee_or_member_of	3683	5589	40537	38068	38580	34847	44339
per:holds_shares_in	0	115	19	21	21	16	20
per:organizations_founded	62	2955	673	642	639	594	741
per:origin	0	0	4089	3872	3826	3783	4170
per:other_family	226	354	469	441	588	444	676
per:parents	876	1615	1297	1246	1228	1083	1390
per:religion	0	6319	134	128	123	116	136
per:schools_attended	185	228	1124	1108	1106	906	1285
per:siblings	446	630	897	874	970	764	1123
per:spouse	1500	2008	2468	2453	2389	2085	2685
per:stateorprovince_of_birth	41	315	36	32	86	89	128
per:stateorprovince_of_death	13	529	55	72	163	151	227
per:statesorprovinces_of_residence	338	1328	2585	2626	14259	12619	30309
per:title	23335	114985	20171	19124	19006	18698	20377
per:top_member_employee_of	205	4277	465	431	4906	4494	5644

Results

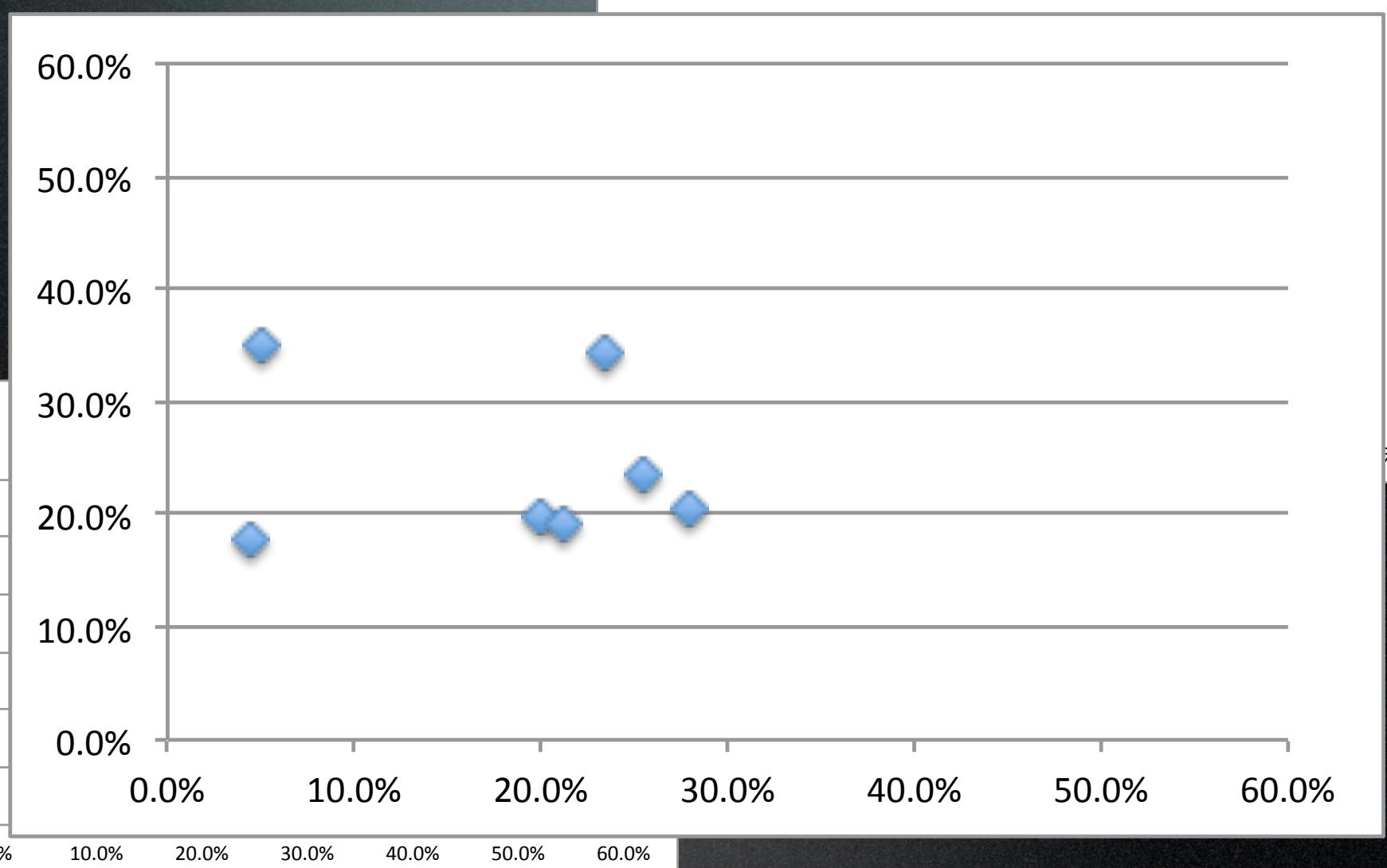


Results

1-Hop Recall/Precision



Results





Ralph Grishman will
now tell us about chasing
squirrels Cold Start
scoring